

Token Emb Temporal Emb



Informer

FC Layer

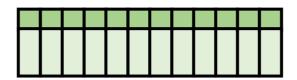
Distil Multi-head ProbSparse Self-attention

Distil

Multi-head ProbSparse Self-attention

**Position Emb** 

Token Emb



**Transformer** 

FC Layer

Multi-head

Self-attention

Multi-head

Self-attention

**Position Emb** 

Simplified Informer

	nheads	num_layers	$e\_layers$	activation	loss	factor	distill
Transformer	4	2	*	*	MSE	*	*
Simplified Informer	4	2	2	$\mathbf{ReLU}$	MAE	2	False
Informer	8	2	3	GELU	MAE	5	True